



# The influence of threat and aversive motivation on conflict processing in the Stroop task

Ming-Ray Liao<sup>1</sup> · Laurent Grégoire<sup>1</sup> · Brian A. Anderson<sup>1</sup>

Published online: 5 July 2020  
© The Psychonomic Society, Inc. 2020

## Abstract

Motivated attention can be driven by the desire to maximize gains or escape punishment. In the Stroop task, when rewards can be obtained by responding quickly to certain colors, corresponding color words are prioritized and produce enhanced interference, suggesting transfer of an attentional bias from color hues to color words. In the present study, we replicated this transfer effect using reward and conducted a parallel experiment exchanging the prospect of reward (appetitive motivation) with the opportunity to avert punishment (aversive motivation). Participants were required to identify the color (hue) of color words and received electric shocks for responses to particular hues that were slow or incorrect. Shock-related words similarly impaired color-ink naming performance. In contrast to prior studies with reward, however, responding to hues associated with shock was also impaired, with threat producing an increase in error rate that ironically resulted in more frequent shocks. Our results suggest that aversive and appetitive motivation affect attention to task-relevant information differently, although each produces a common bias in automatic stimulus processing presumably driven by valence.

**Keywords** Stroop · Reinforcement learning · Motivation · Punishment · Aversive conditioning

Punishment, along with reward and other extrinsic motivators, are an important determinant of human and animal behavior (Maren, 2001; Schultz, 2004; Thorndike, 1911). Threat detection is especially important to prevent and avoid harm, and aversive outcomes have been posited to heighten sensitivity to negative information (Vuilleumier, 2005, for a review). Reward has been shown to facilitate both visual search and cognitive control, possibly through increased neural representation of reward-associated stimuli via biased competition (Anderson, 2019; Bijleveld, Custers, & Aarts, 2010;

Knutson, Adams, Fong, & Hommer, 2001). Consequently, task-irrelevant reward-associated stimuli can impair task performance (Anderson, 2016, for a review). Some studies have shown punishment to have effects similar to reward (Anderson & Britton, 2019; H. Kim & Anderson, 2020; Nissens, Failing, & Theeuwes, 2017; Schmidt, Belopolsky, & Theeuwes, 2015, 2017; Van Damme et al., 2004). For example, Van Damme et al. (2004) reported that visual targets were detected faster when their location was predicted by a cue conditioned with a pain signal, relative to a neutral cue. When presented as a distractor in a visual search task, a stimulus previously conditioned with an aversive electrical stimulation impairs performance compared with a neutral stimulus, independent of perceptual salience (Schmidt et al., 2015). Thus, the effect of punishment on information processing can be conceptualized as being similar to reward, where both types of outcome serve as motivational inputs to attention and cognition (Engelmann & Pessoa, 2007). How extrinsic motivators shape behavior in the presence of conflicting information, as in the Stroop task or other conflict paradigms, is less understood.

The classic Stroop task entails the presentation of stimuli comprising two dimensions, word and color (Stroop, 1935; see MacLeod, 1991, for a review). Participants are required

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.3758/s13414-020-02072-5>) contains supplementary material, which is available to authorized users.

---

✉ Ming-Ray Liao  
m4liao@tamu.edu

Laurent Grégoire  
lgregoire1@tamu.edu

Brian A. Anderson  
brain.anderson@tamu.edu

<sup>1</sup> Department of Psychological & Brain Sciences, Texas A&M University, 4235 TAMU, College Station, TX 77843-4235, USA

to identify the ink color (hue) of color words while ignoring their meaning. Performance is usually impaired when the two dimensions are incongruent (e.g., the word “red” presented in blue ink) and improved when the two dimensions are congruent (e.g., the word “blue” presented in blue ink), relative to a neutral situation (e.g., a nonword presented in blue ink). It has been proposed that automatic reading of the irrelevant dimension (i.e., word) interferes with processing of the relevant dimension (i.e., ink color; Carter & Van Veen, 2007; Cohen, Dunbar, & McClelland, 1990). Krebs, Boehler, and Woldorff (2010) investigated the effect of appetitive motivation in the Stroop task by associating two ink colors with the opportunity to earn a monetary reward for fast and accurate responses, and found improvements in color naming for reward-associated hues. They also found that there were greater performance deficits for reward-related words (appetitive association) compared with words unrelated to reward, suggesting transfer of a motivated attentional bias from one feature dimension (hue) to another (word meaning).

Prospect theory proposes that humans are more sensitive to losses than to reward (Kahneman & Tversky, 1979; but see Ert & Erev, 2013; Lejarraga, Schulte-Mecklenbeck, Pachur, & Hertwig, 2019; Yechiam, 2019; Yechiam & Hochman, 2013a, 2013b). Loss aversion explains why people are more willing to take risks to avoid a loss than to procure a gain (Schindler & Pfattheicher, 2017), and how penalties are more effective than reward in motivating people (Gächter, Orzen, Renner, & Starmer, 2009). However, Wentura, Müller, and Rothermund (2014) showed that there were no differences in the magnitude of attentional capture by reward-associated and loss-associated distractors (see also Grégoire, Britton, & Anderson, 2020). Carsten, Hoofs, Boehler, and Krebs (2019) tested predictions arising from prospect theory by modifying the paradigm from Krebs et al. (2010) to compare how rewards and losses facilitated task performance by having one hue associated with reward and the other associated with loss. While both rewards and losses improved Stroop task performance, participants were slower in loss trials compared with reward trials. The authors suggested that the valence of the association may have modulated responses such that positive valence facilitated approach behavior, while negative valence had a contradictory mapping with avoidance behavior that had to be overcome (Guitart-Masip, Duzel, Dolan, & Dayan, 2014). In other words, loss is potentially inducing both appetitive and aversive associations, which differently affect attention.

The present study follows up on Krebs et al.’s (2010) and Carsten et al.’s (2019) studies, but we exchanged the prospect of reward with the opportunity to avert punishment as administered with electric shocks. Monetary loss may have been more motivating with respect to negative reinforcement, especially in the context of other available gains, leading to a focus on maximizing overall value and the results seen in Carsten

et al. (2019). We therefore opted for a punisher that is inherently aversive and would serve as a more potent unconditioned stimulus in the context of associative learning. In a color-word Stroop task, two ink colors were associated with shock, while two other ink colors were never followed by shock. Participants were asked to identify the ink color, ignoring the semantic meaning of the word. For ink colors associated with shock, slow responses and errors resulted in the delivery of shock. Much like rewarding fast responses, this manipulation provided (in this case, aversive) motivation to maximize performance on trials with a shock-associated ink color. Although prior research has investigated how the emotional content of the irrelevant dimension interferes with task performance (Algom, Chajut, & Lev, 2004; Fox, Russo, Bowles, & Dutton, 2001; Fox, Russo, & Dutton, 2002; Stormark, Nordby, & Hugdahl, 1995; Yiend & Mathews, 2001), few studies have applied that information to the relevant dimension (color hue) via associative learning.

If the prospect of procuring a reward, averting a loss, and averting an aversive physical punisher have similar motivational influences on attention, we should observe the same effects that Krebs et al. (2010) and Carsten et al. (2019) did—a general facilitation of performance (faster responses and increased accuracy) for shock-associated colors. In addition, the attentional bias for shock-associated colors should spill over to corresponding words, impeding performance on incongruent trials with shock-related words. Alternatively, punishment associations might have a generally detrimental effect on performance, similar to emotionally charged words (e.g. Algom et al., 2004; Williams, Mathews, & MacLeod, 1996).

## Experiment 1

### Method

#### Participants

Thirty-six participants, between the ages of 18 and 35 years inclusive, were recruited from the Texas A&M University community. All reported normal or corrected-to-normal visual acuity, normal color vision, and provided written informed consent. Data from two participants were excluded from the analyses due to low proportion of correct responses. Specifically, the error and miss rates exceeded three standard deviations from the group mean. The final sample included 34 participants (24 females), with a mean age of 19.62 years ( $SD = 1.92$ ). Using the smallest effect size for the main effect of reward on either the relevant (hue) or irrelevant (word) stimulus dimension in Krebs et al. (2010) yielded power  $\beta > 0.90$  with  $\alpha = 0.05$  for the obtained sample size (G\*Power 3.1). Participants were compensated with US\$10 or course credit.

All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

**Apparatus**

A Dell OptiPlex equipped with the MATLAB software and Psychophysics Toolbox extensions (Brainard, 1997) was used to present the stimuli on a Dell P2717H monitor. Electric shocks were administered to participants’ left forearms via paired electrodes (EL500, BioPac Systems, Goleta, CA, USA) linked to an isolated linear stimulator (STMISOLA, BioPac Systems) set to constant current and controlled by MATLAB scripts.

**Stimuli**

The stimuli consisted of five color words: *red*, *green*, *blue*, *yellow*, and *brown*. The words appeared in lowercase font (Arial 40) and were presented in four prototypical ink colors: red, green, blue, and yellow. The semantic meaning of a given word could be congruent (e.g. “red” written in red ink) or incongruent (e.g. “red” written in green ink) with regard to the ink color. Furthermore, trials consisting of words with no response mapping (e.g. “brown” written in red ink) were intermixed to provide a neutral category (see Fig. 1). It is important to note that the word “brown” is neutral with respect to valence in that it is neither associated with shock nor the absence of shock (safe response) as a reported color, although it is not neutral with respect to Stroop congruity (even though it does not have an explicit response mapping, the word is still incongruent with the ink color). It is in this sense of valence

that we use the label “neutral,” consistent with Krebs et al. (2010). Of interest in the present study is the comparison of words and ink colors that differ with respect to their relationship to shock, and in this context “neutral” words merely provide an opportunity to subtract out any effect of the ink color on responses when assessing the influence of the valence of the words. A complete breakdown of how all of the different color/word combinations differ with respect to outcomes (shock), Stroop congruity, and response mapping is provided in supplemental Table S1.

**Procedure**

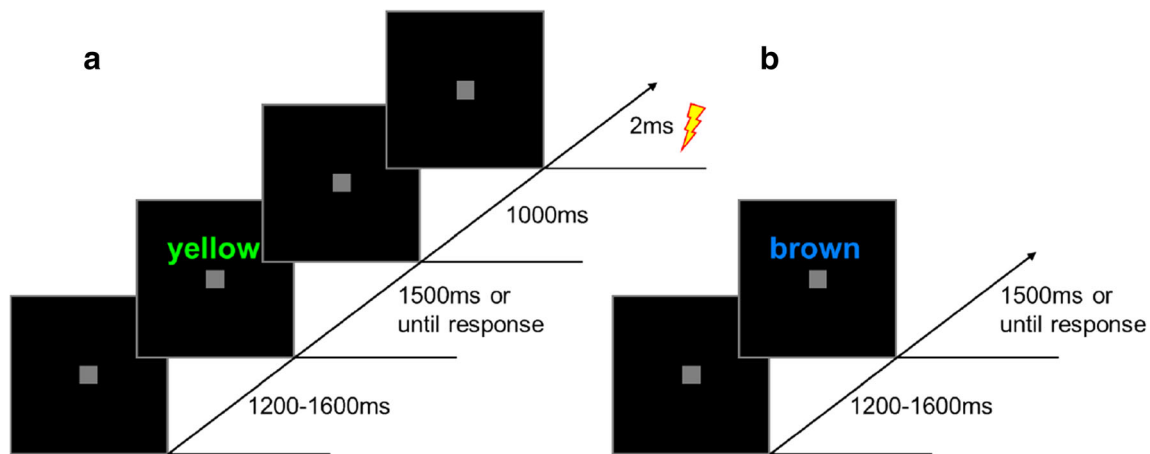
Throughout the experiment, a gray fixation square (visual angle 0.5°) was maintained at the center of a black background screen. Each trial began with the fixation square for a duration that varied randomly between 1,200 and 1,600 ms, followed by the Stroop stimulus (vertical visual angle between 1.3° and 1.5°, horizontal visual angle between 2.6° and 5.6°) on top of the square for 1,500 ms or until response. On punished trials, a 2-ms shock was administered 1,000 ms after the Stroop stimulus (see Fig. 2).

Participants were required to report the ink color of words as fast and accurately as possible while ignoring their semantic meaning. Participants were instructed to manually enter their responses using a standard (QWERTY layout) keyboard with their left and right index and middle fingers on the “z,” “x,” “n,” and “m” keys. Each key was mapped onto one of four colors. Color-key assignments were pseudorandomized across participants such that there would always be one color associated with shock for each hand. This was to avoid participants associating one hand with punishment.

		Word (irrelevant)			
		congruent	incongruent		neutral
Color (relevant)	Potential-shock				
	No-shock				

**Fig. 1** Stimuli and experimental conditions (counterbalanced across participants). A subset of ink colors was associated with the potential of shock (potential shock; e.g., red and green), while the remaining ink colors were not (no shock; e.g., blue and yellow). The word meaning

(irrelevant dimension) could be congruent, incongruent shock unrelated, incongruent shock related, or neutral with regard to the valence of the ink color. (Color figure online)



**Fig. 2** Sequence of events in the color-Stroop task. The fixation square remained in the middle of the screen for the entire task. Each trial began with the fixation square for a random duration between 1,200 and 1,600 ms, followed by the stimulus on top of the square for 1,500 ms or until

response. When a shock is delivered as in **a**, the fixation square remained for 1,000 ms before the shock is administered for 2 ms. Otherwise, the trial ends as in **b**. (Color figure online)

Before the Stroop task, participants had eight trials to learn the color-key pairing with colored squares (visual angle  $3.0^\circ$ ). A colored square remained on the screen until the correct key for that color was pressed. Next, participants completed eight practice trials of the Stroop task with no time limit and feedback messages if they pressed an incorrect key, after which there were forty more practice trials with no feedback and a 1,500 ms time limit. Participants were then connected to the shock device, and shock intensity was adjusted individually to be “unpleasant but not painful” (Grégoire & Greening, 2019, 2020; A. J. Kim & Anderson, 2019; Murty, LaBar, & Alison Adcock, 2012; Schmidt et al., 2015, 2017). Finally, participants completed four runs of the Stroop task: two runs with shocks (learning phase) followed by two runs without shock (extinction phase). Each run consisted of 120 trials for a total of 480 trials.

The 120 trials per run consisted of six blocks of 20 randomized trials. Each block consisted of all the possible color-word and color-ink pairings, such that there were four congruent trials and 12 incongruent trials, and four neutral trials. Half of each block (10 trials) were with shock-associated ink. The response time (RT) cutoff was adjusted dynamically to be the median of the participants’ RT over the past 40 trials. If the participant responded incorrectly or slower than the cutoff on trials with shock-associated ink, then participants would receive an electric shock, with similar performance-dependent frequency as in Krebs et al. (2010). There were no consequences for responding incorrectly or slowly for non-shock-associated ink colors. However, there was a six-shock limit per block. Before performing the learning phase, participants were informed that they might receive a shock “if [they] respond slowly or incorrectly.” Before performing the extinction phase, participants were explicitly instructed that no more electric stimulation would be delivered, and shock electrodes were disconnected.

## Data analysis

Analyses were conducted separately for learning phase (i.e., the two runs with shock) and extinction phase (i.e., the two runs without shock). RTs for correct responses beyond three standard deviations of the conditional mean (<1% of responses for all analyses) were removed from the data for each participant (Grégoire, Perruchet, & Poulin-Charronnat, 2013, 2014, 2015). The remaining data were submitted to repeated-measure analysis of variance (ANOVA), with the aversive association (color: potential-shock, no-shock) and the congruity (word: congruent, incongruent, [valence] neutral) as within-subjects variables. The same analysis was conducted for error rates, which included misses and incorrect responses. In order to investigate differential effects of *shock-related* and *shock-unrelated* word information, additional ANOVAs were conducted focusing on incongruent trials, with color ink (color: potential shock, no shock) and color word (shock related, shock unrelated) as within-subjects variables. The error rates and RTs used for these analyses were relative to the corresponding neutral condition (i.e., performance from the neutral condition with the same shock-ink association was subtracted; see Krebs et al., 2010) in order to more effectively isolate the influence of the words’ association with shock.

## Results

### Stroop effect

The difference in errors and RT in the learning phase amounted to a Stroop effect (incongruent – congruent) of 3.25%,  $t(33) = 5.18$ ,  $p < .001$ ,  $d = 0.89$ , and 51 ms,  $t(33) = 8.55$ ,  $p < .001$ ,  $d = 1.47$ . The difference in RT in the extinction phase amounted to a Stroop effect of 42 ms,

$t(33) = 5.53, p < .001, d = 0.95$ , but there was no Stroop effect for errors ( $p = .124$ ).

### Learning phase

**Errors** The ANOVA revealed a significant main effect of aversive association,  $F(1, 33) = 4.68, p = .038, \eta_p^2 = 0.124$ , with more errors in potential-shock compared with no-shock trials. We also observed a significant main effect of congruity,  $F(2, 66) = 11.65, p < .001, \eta_p^2 = 0.261$ , such that errors varied as a function of condition—incongruent, congruent, and neutral (see Fig. 3a & Table 1). Errors were more frequent for incongruent compared with congruent and neutral conditions ( $ts > 3.25, ps < .004, ds > 0.55$ ). The interaction between aversive association and congruity was not significant,  $F(2, 66) = 0.53, p = .590$ .

Focusing on incongruent trials only, we observed a main effect of color word,  $F(1, 33) = 7.12, p = .012, \eta_p^2 = 0.177$ , with more errors for shock-related words compared with shock-unrelated words (see Fig. 3b). The main effect for the color-ink was not significant,  $F(1, 33) = 0.70, p = .409$ . We also observed a significant interaction between color word and color ink,  $F(1, 33) = 4.16, p = .049, \eta_p^2 = 0.112$ . Errors were comparatively more frequent on incongruent trials with shock-related words and no-shock ink compared with incongruent trials with shock-unrelated words, which was confirmed by pairwise comparisons ( $ts > 2.59, ps < .015, ds > 0.44$ ).

**Response times** An ANOVA revealed a significant main effect of congruity,  $F(2, 66) = 35.77, p < .001, \eta_p^2 = 0.520$ . Incongruent words were slower than congruent and neutral words, and neutral words were slower than congruent words ( $ts > 3.28, ps < .003, ds > 0.56$ ; see Fig. 4a). The main effect of

aversive association was not significant, nor was the interaction between aversive association and congruity ( $Fs < 1$ ).

Focusing on incongruent trials, there was a main effect of color word,  $F(1, 33) = 7.54, p = .010, \eta_p^2 = 0.186$ , such that shock-related words slowed responses more than shock-unrelated words (see Fig. 4b). There was no main effect of color ink and no interaction effect ( $Fs < 1$ ).

### Extinction phase

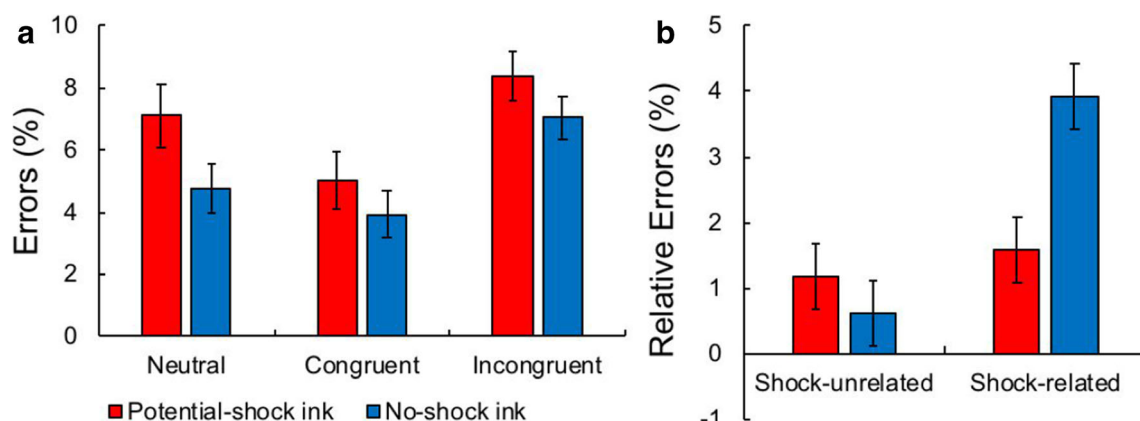
**Errors** An ANOVA with aversive association (color: potential shock, no shock) and congruity (word: congruent, incongruent, neutral) as within-subjects variables revealed no significant main effect or interaction (all  $ps > 0.10$ ; see Table 1). The ANOVA on incongruent trials yielded no significant main effect or interaction (all  $ps > 0.10$ ).

**Response times** An ANOVA revealed a significant main effect of congruity,  $F(2, 66) = 16.08, p < .001, \eta_p^2 = 0.328$ . Incongruent words were slower than congruent and neutral words, and neutral words were slower than congruent words ( $ts > 2.44, ps < 0.021, ds > 0.41$ ; see Table 1). There was no main effect of aversive association,  $F(1, 33) = 0.64, p = .430$ , and a marginally significant interaction effect,  $F(2, 66) = 2.80, p = .068, \eta_p^2 = 0.078$ .

Focusing on incongruent trials, there was a marginal effect of color word that mirrored the pattern during the training phase,  $F(1, 33) = 3.84, p = .059, \eta_p^2 = 0.104$ . The main effect of color ink and its interaction with color word were not significant (all  $ps > 0.10$ ).

### Discussion

In Experiment 1, participants performed a modified version of the color-word Stroop task in which a subset of ink colors



**Fig. 3** Mean error rates for the learning phase. **a** Error rates between potential-shock (red bars) and no-shock ink (blue bars) for congruent, neutral, and incongruent trials. Error bars represent standard error of the

means. **b** Error rates for shock-related and shock-unrelated incongruent trials are depicted as the difference relative to corresponding neutral words. (Color figure online)



**Table 1** Effects of relevant and irrelevant shock associations on performance

Color (aversive association)		Word (congruity)			
		Congruent	Incongruent		Neutral
	Shock unrelated		Shock related		
<b>Learning phase</b>					
Potential shock	RT (ms)	651.16 (94.25)	694.07 (85.60)	718.09 (107.35)	684.33 (91.81)
	Error rate	5.02 (5.50)	8.29 (4.79)	8.70 (6.52)	7.11 (6.03)
No shock	RT (ms)	642.18 (90.38)	692.00 (80.45)	702.07 (89.13)	674.26 (80.37)
	Error rate	3.92 (4.46)	5.39 (4.76)	8.70 (4.96)	4.78 (4.60)
<b>Extinction phase</b>					
Former potential shock	RT (ms)	633.98 (75.30)	680.54 (66.13)	690.80 (65.75)	668.11 (74.85)
	Error rate	5.99 (4.94)	5.69 (4.10)	7.16 (5.92)	5.21 (4.01)
Former no shock	RT (ms)	647.65 (75.88)	670.68 (66.97)	685.13 (71.28)	649.33 (58.95)
	Error rate	4.56 (5.31)	6.58 (5.35)	7.75 (4.51)	5.34 (5.83)

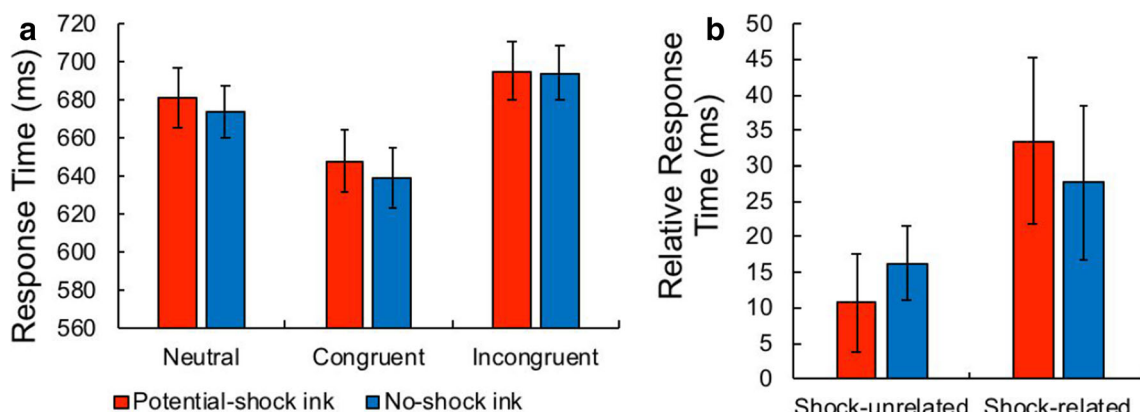
Note. Standard deviations are in parentheses

were associated with punishment (shocks). Participants had an opportunity to avoid shock if they responded accurately and fast enough. In contrast with prior research using reward (Bijleveld et al., 2010; Engelmann & Pessoa, 2007; Krebs et al., 2010) and loss (Müller, Rothermund, & Wentura, 2016; Wentura et al., 2014) as a motivator, shock association (in both color ink and color words) interfered with, rather than enhanced, task performance.

In the learning phase, there was an overall increase in errors for trials with potential-shock ink compared with no-shock ink, as well as an overall increase in both errors and RT for trials with shock-related words compared with shock-unrelated words. The difference in RT remained marginally significant in extinction, while the difference in errors did not. The observed difference in error rate was particularly striking, as participants made the most errors on trials on which they

should have been the most motivated to perform well (to avoid shock), which was distinctly counterproductive in that it resulted in more overall shocks being delivered. That is, not only did aversive motivation fail to improve task performance, it ironically impaired performance. Although the typical Stroop effect was observed across all incongruent trials, such interference was particularly pronounced for shock-related words, consistent with prior results investigating the consequences of color–reward associations (Krebs et al., 2010).

Because our design involved more incongruent trials compared with congruent trials overall (as in Krebs et al., 2010), there was inevitably more shocks received on incongruent trials. Our results could have been influenced by this imbalance in design, which may have affected baseline performance on incongruent trials, but it cannot account for the differential processing of shock-associated stimuli. Such an imbalance is



**Fig. 4** Mean RT from the learning phase. **a** RTs between potential-shock (red bars) and no-shock ink (blue bars) for congruent, neutral, and incongruent trials. Error bars represent standard error of the means. **b** RTs for

shock-related and shock-unrelated incongruent trials are depicted as the difference relative to corresponding neutral words. (Color figure online)

necessary to avoid biases in how frequently each color-word combination is presented, which could compromise the comparisons of interest.

A pronounced difference in our design compared with that of Krebs et al. (2010) is that participants were not explicitly informed of the colors associated with shocks. Awareness may have contributed to more efficient deployment of attention to the relevant colors to maximize gains and minimize loss (Carsten et al., 2019; Krebs et al., 2010). Implicit learning of associations may also have contributed to the persistence of attentional biases in RT. In order to draw direct contrasts between aversive and appetitive motivation in the Stroop task, we conducted a second experiment that exchanged the opportunity to avoid shocks with the prospect to earn rewards while keeping everything else the same. In order to better understand the potential role of awareness of the task contingencies, we also included an awareness assessment at the end of the experiment.

## Experiment 2

### Method

#### Participants

Thirty-six new participants, between the ages of 18 and 35 years inclusive, were recruited from the Texas A&M University community. All reported normal or corrected-to-normal visual acuity, normal color vision, and provided written informed consent. Data from one participant was excluded from the analyses due to low proportion of correct responses. Specifically, the error and miss rates exceeded three standard deviations from the group mean. The final sample included 35 participants (13 females), with a mean age of 21.97 years ( $SD = 3.05$ ). Participants were compensated with earnings from the task. All procedures were approved by the Texas A&M University Institutional Review Board and conformed with the principles outlined in the Declaration of Helsinki.

#### Apparatus, stimuli, procedure, and data analysis

Apparatus remained the same as in Experiment 1, but without the use of an isolated linear stimulator since no electric shocks were involved. Stimuli and procedure remained largely the same, but the opportunity to avoid shocks was replaced with the prospect for reward. Participants were informed that “the rewards will be more likely if [they] respond fast and correctly,” to mirror the instructions given in Experiment 1. On trials with reward-associated ink, participants could obtain 20¢ rewards if they responded correctly and faster than the RT cutoff. Reward feedback was displayed on every trial after the stimulus display and consisted of the amount earned on the trial (+20¢ or +0¢), as well as the total amount earned. The RT

cutoff was dynamically adjusted to be the median of the participants’ RT over the past 40 trials, as in Experiment 1. The six-shock limit per block was replaced with a six-reward limit per block, meaning that the maximum amount participants could make totaled to \$14.40. Before performing the extinction phase, participants were explicitly instructed that no more rewards would be given.

At the end of the experiment, participants were asked to rate how likely they were to receive a reward for the displayed stimulus by clicking on a number line that spanned from 0 to 100. Each color and word combination were shown once, resulting in 20 trials. For each participant, the average score for no-reward ink was subtracted from the average score for potential-reward ink, resulting in an awareness score with positive scores reflecting reward expectancies that were in the direction of the actual contingencies used in the learning phase, and with larger scores reflecting more pronounced awareness. With respect to RT and error rate, the same analyses as in Experiment 1 were conducted.

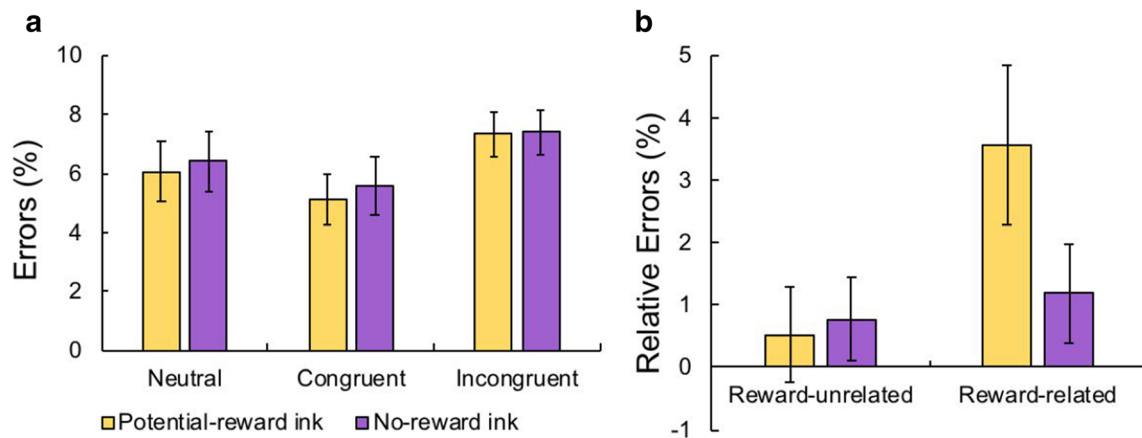
### Results

#### Stroop effect

The difference in errors and RT between congruent and incongruent trials in the learning phase amounted to a Stroop effect of 2.02%,  $t(34) = 2.87$ ,  $p = .007$ ,  $d = 0.49$ , and 52 ms,  $t(34) = 6.74$ ,  $p < .001$ ,  $d = 1.14$ . The difference in RT in the extinction phase amounted to a Stroop effect of 51 ms,  $t(34) = 9.26$ ,  $p < .001$ ,  $d = 1.57$ , but there was no Stroop effect in errors ( $p = .208$ ).

#### Learning phase

**Errors** An ANOVA revealed a significant main effect of congruity,  $F(1, 34) = 4.40$ ,  $p = .016$ ,  $\eta_p^2 = 0.115$ , with more frequent errors for incongruent compared with neutral and congruent conditions ( $ts > 2.20$ ,  $ps < .030$ ,  $ds > 0.37$ ). The main effect of appetitive association and the interaction effect were not significant ( $F_s < 1$ ; see Fig. 5a & Table 2). Focusing on incongruent trials only, we observed a main effect of color word,  $F(1, 34) = 6.00$ ,  $p = .020$ ,  $\eta_p^2 = 0.150$ , with more frequent errors for reward-related compared with reward-unrelated words (see Fig. 5b). The main effect of color ink was not significant,  $F(1, 34) = 0.93$ ,  $p = .341$ . The interaction effect between color word and color ink was marginally significant,  $F(1, 34) = 4.07$ ,  $p = .052$ ,  $\eta_p^2 = 0.107$ . Errors were comparatively more frequent on incongruent trials with potential-reward ink and reward-related words compared with potential-reward ink and reward-unrelated words.



**Fig. 5** Mean error rates for the learning phase. **a** Error rates between potential-reward (yellow bars) and no-reward ink (purple bars) for congruent, neutral and incongruent trials. Error bars represent standard error

of the means. **b** Error rates for shock-related and shock-unrelated incongruent trials are depicted as the difference relative to corresponding neutral words. (Color figure online)

**Response times** An ANOVA revealed a significant main effect of congruity,  $F(1, 34) = 34.41, p < .001, \eta_p^2 = 0.503$ . Participants were slower at responding to incongruent trials than congruent and neutral trials, and slower at neutral trials compared with congruent trials ( $t_s > 3.43, p_s < .002, d_s > 0.58$ ; see Fig. 6a). The main effect of appetitive association and its interaction with congruity were not significant ( $F_s < 1$ ). The ANOVA on incongruent trials yielded a significant main effect of color word only,  $F(1, 34) = 12.13, p = .001, \eta_p^2 = 0.263$ , with slower responses for reward-related words compared with reward-unrelated words (see Fig. 6b). There was no main effect for color ink or interaction between color ink and color word ( $F_s < 1$ ).

**Extinction phase**

**Errors** An ANOVA with appetitive association (color: potential reward, no reward) and congruity (word: congruent, incongruent, neutral) as within-subjects variables revealed no significant main effect or interaction (all  $p_s > .28$ ; see Table 2). The ANOVA on incongruent trials yielded a significant main effect of color word,  $F(1, 34) = 6.00, p = .020, \eta_p^2 = 0.150$ , with more errors for reward-related words compared with reward-unrelated words, but no main effect of color ink or interaction ( $F_s < 1$ ).

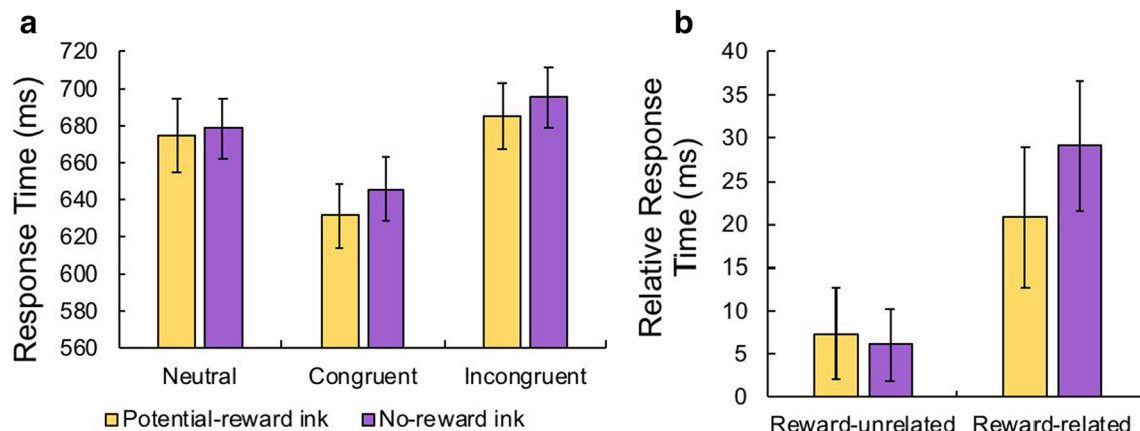
**Response times** An ANOVA revealed a main effect of congruity,  $F(1, 34) = 42.48, p < .001, \eta_p^2 = 0.555$ . Participants were

**Table 2** Effects of relevant and irrelevant reward associations on performance

Color (appetitive association)		Word (congruity)			
		Congruent	Incongruent		Neutral
			Reward unrelated	Reward related	
<b>Learning phase</b>					
Potential reward	RT (ms)	631.52 (103.12)	682.12 (109.43)	695.63 (100.12)	674.72 (119.55)
	Error rate	5.12 (4.96)	6.59 (5.11)	9.64 (6.13)	6.07 (5.92)
No reward	RT (ms)	645.63 (102.37)	684.76 (96.03)	707.79 (101.21)	678.64 (94.67)
	Error rate	5.60 (5.89)	7.20 (5.09)	7.62 (5.03)	6.43 (6.09)
<b>Extinction phase</b>					
Former potential reward	RT (ms)	610.11 (88.47)	663.82 (81.02)	683.74 (94.07)	650.92 (74.31)
	Error rate	6.43 (6.73)	7.22 (4.63)	9.29 (6.48)	6.91 (5.34)
Former no reward	RT (ms)	615.52 (87.17)	662.33 (83.59)	673.82 (94.55)	661.94 (92.97)
	Error rate	7.62 (8.18)	7.50 (6.17)	8.33 (5.15)	6.79 (7.70)

Note. Standard deviations are in parentheses



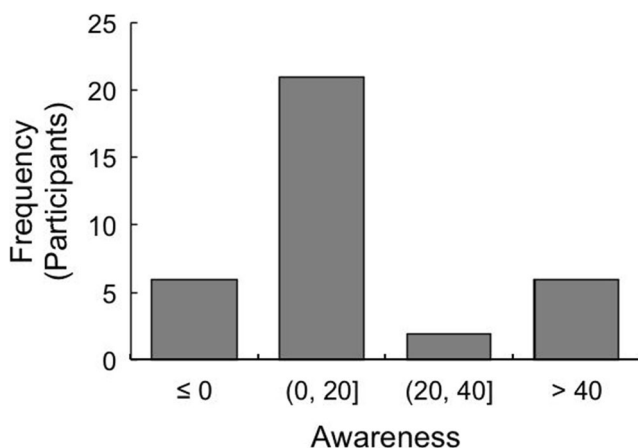


**Fig. 6** Mean RT from the learning phase. **a** RTs between potential-reward (yellow bars) and no-reward ink (purple bars) for congruent, neutral, and incongruent trials. Error bars represent standard error of the

means. **b** RTs for reward-related and reward-unrelated incongruent trials are depicted as the difference relative to corresponding neutral words. (Color figure online)

slower on incongruent trials compared with neutral and congruent trials, and slower at neutral trials compared with congruent trials ( $t_s > 2.07$ ,  $p_s < 0.045$ ,  $d_s > 0.34$ , see Table 2). There were no main effect of appetitive association or interaction effect ( $F_s < 1$ ). An ANOVA on incongruent trials yielded a main effect of color word,  $F(1, 34) = 7.33$ ,  $p = .011$ ,  $\eta_p^2 = 0.177$ , a marginal main effect for color ink,  $F(1, 34) = 3.63$ ,  $p = .065$ ,  $\eta_p^2 = 0.097$ , but no significant interaction effect,  $F(1, 34) = 0.48$ ,  $p = .492$ .

**Awareness check** Participants had an average awareness score of 15.4 ( $SD = 26.3$ ). There was a correlation between awareness and the difference in RT between potential-reward and no-reward trials during the learning phase,  $r(33) = .49$ ,  $p = .003$ , but not for errors,  $r(33) = .28$ ,  $p = .104$ . One participant had an RT difference greater than three standard deviations from the mean; with this individual excluded from the analyses, the correlations were no longer significant, ( $p_s > .150$ ) (Fig. 7).



**Fig. 7** Histogram of the awareness check where the higher the awareness score, the more participants were aware of which color inks were associated with reward

## Discussion

In Experiment 2, we replicate the cost associated with reward-related words observed in Krebs et al. (2010) and Carsten et al. (2019), mirroring the spillover consequence of aversive motivation observed in Experiment 1. Unlike in Krebs et al. (2010), this effect persisted into an extinction phase. We did not observe a cost associated with ink colors signaling potential reward, in contrast to Experiment 1, with the direction of effect being opposite that observed with aversive motivation. At the same time, we do not see the performance benefit on reward-associated ink-color trials observed in Krebs et al. (2010), which might be explained by the generally low awareness of the reward contingencies in our study, as further suggested by the relationship between awareness and task performance. In this sense, our findings additionally demonstrate that the spillover effect observed for reward-related words is more likely the consequence of associative learning rather than a peripheral consequence of explicit strategy, consistent with the principle of cross-dimension value-driven attention (Anderson, 2015, 2016; Anderson, Laurent, & Yantis, 2011, 2012) and the finding that value-based interference can be observed in participants unaware of the reward contingencies (Grégoire & Anderson, 2019).

## General discussion

Prior studies have investigated how reward and monetary loss facilitate performance in cognitive control and attention tasks (Bijleveld et al., 2010; Carsten et al., 2019; Engelmann & Pessoa, 2007; Krebs et al., 2010; Müller et al., 2016; Wentura et al., 2014). In Experiment 1 of the present study, our aversive conditioning procedure encouraged participants

to improve task performance on certain color-ink trials. However, we observed an impairment on these trials instead.

This result is surprising, given that past studies might have suggested otherwise, with punishment biasing attention in a similar manner to reward (Grégoire et al., 2020; H. Kim & Anderson, 2020; Nissens et al., 2017; Schmidt et al., 2015, 2017; Van Damme et al., 2004). However, this result can be better understood under the approach-avoidance framework (Aupperle, Melrose, Francisco, Paulus, & Stein, 2015; Guitart-Masip et al., 2014). Krypotos and colleagues used an approach-avoidance task (AAT) and found that participants were slower to approach shock-associated stimuli and faster at avoiding them (Krypotos, Effting, Arnaudova, Kindt, & Beckers, 2014). What is happening in our task is similar in that responses are tied to both positive and negative outcomes—participants must make a response to avoid shock, but in responding might cause the shock to be delivered sooner. The color-word Stroop task was not designed to dissociate approach and avoidance behavior (see Chajut, Mama, Levy, & Algom, 2010). However, this allows us to dissociate motivation from valence.

With reward contingencies, distraction effects from reward-related words (Carsten et al., 2019; Krebs et al., 2010) could be explained as a spillover consequence of motivated attention (see Anderson, 2018): Participants preferentially attended to reward-associated colors in an effort to maximize gains, and this bias could not be restricted to the task-relevant (color-ink) dimension. This is line with theories that suggest reward increases biased competition, essentially improving the signal-to-noise ratio of reward-associated stimuli (Della Libera & Chelazzi, 2006; Desimone & Duncan, 1995; Engelmann & Pessoa, 2007). Although we do not see a benefit for reporting colors associated with punishment in the present study, we do see an attentional bias for shock-related words that impairs performance on incongruent trials, consistent with prior studies using monetary rewards and losses (Carsten et al., 2019; Krebs et al., 2010) that we replicate here. Our results therefore imply that the influence of reward and punishment on incongruent trials reflects a valence-dependent bias in information processing that is more than an extension of motivated, goal-directed attention; that is, stimuli associated with a valent outcome such as reward or punishment can capture attention when appearing in the task-irrelevant dimension (color word), regardless of whether such stimuli are actively prioritized in the task-relevant dimension (color ink). In contrast, aversive and appetitive motivation have distinct influences on the processing of task-relevant information, with only aversive motivation impairing performance when the task involves responding quickly and accurately in the face of conflict.

In Krebs et al. (2010), reward-associated words ceased to affect performance in an extinction phase, while we observed persistent influences for both reward and punishment

associations. Explicit awareness of contingencies allows participants to engage and disengage their attentional set towards relevant colors to maximize gains and minimize losses (Carsten et al., 2019), while also allowing participants to relax their constraints when they know that nothing is at stake. This stands in contrast to studies where participants did not have explicit information concerning the task contingencies and the attentional biases persists for an extended period of time (Anderson & Britton, 2019; Anderson et al., 2011; Anderson & Yantis, 2013; Anderson et al., 2016; Nissens et al., 2017; Schmidt et al., 2015, 2017; Van Damme et al., 2004). It is possible that appetitive motivation is an effective facilitator of goal-directed attentional processes particularly when participants know what the contingencies are (Krebs et al., 2010) and can strategically prioritize information processing accordingly, which was not the case in our experiments. In contrast, the automatic component of the color-outcome associations, as reflected in interference from color words, results from associative learning and is nonstrategic, as evidenced by the robust influence with limited awareness observed in the present study. Explicit awareness of the contingencies may facilitate more rapid extinction when the valent outcomes are removed from the task, which might explain the persistence we observed in our study.

Future investigations could directly compare the effect of awareness on how reward, loss, and shock associations benefit and impair performance in conflict paradigms. Another interesting investigation could be to directly compare emotionally charged stimuli and stimuli that have been paired with punishment in the context of a Stroop task. The distracting effects are phenomenologically similar, but there may be differences between naturally negative associations and negatively conditioned associations. Lastly, the differences in how primary and secondary aversive motivations affect behavior is still an open question.

**Open practices statement** The reported experiment was not preregistered. Raw data for the experiment are downloadable as supplemental material.

**Acknowledgements** This study was supported by grants from the Brain & Behavior Research Foundation [NARSAD 26008] and NIH [R01-DA046410] to B.A.A. The authors declare no competing interests.

## References

- Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General*, 133(3), 323–338. doi:<https://doi.org/10.1037/0096-3445.133.3.323>
- Anderson, B. A. (2015). Value-driven attentional priority is context specific. *Psychonomic Bulletin & Review*, 22(3), 750–756. doi:<https://doi.org/10.3758/s13423-014-0724-0>

- Anderson, B. A. (2016). The attention habit: how reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, 1369(1), 24–39. doi:<https://doi.org/10.1111/nyas.12957>
- Anderson, B. A., Folk, C. L., Garrison, R., & Rogers, L. (2016). Mechanisms of habitual approach: Failure to suppress irrelevant responses evoked by previously reward-associated stimuli. *Journal of Experimental Psychology: General*, 145(6), 796–805. doi:<https://doi.org/10.1037/xge0000169>
- Anderson, B. A. (2018). Controlled information processing, automaticity, and the burden of proof. *Psychonomic Bulletin & Review*, 25, 1814–1823. doi:<https://doi.org/10.3758/s13423-017-1412-7>
- Anderson, B. A. (2019). Neurobiology of value-driven attention. *Current Opinion in Psychology*, 29, 27–33. doi:<https://doi.org/10.1016/j.copsyc.2018.11.004>
- Anderson, B. A., & Britton, M. K. (2019). On the automaticity of attentional orienting to threatening stimuli. *Emotion*. doi:<https://doi.org/10.1037/emo0000596>
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, 108(25), 10367–10371. doi:<https://doi.org/10.1073/pnas.1104047108>
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2012). Generalization of value-based attentional priority. *Visual Cognition*, 20(6), 647–658. doi:<https://doi.org/10.1080/13506285.2012.679711>
- Anderson, B. A., & Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, 39(1), 6–9. doi:<https://doi.org/10.1037/a0030860>
- Aupperle, R. L., Melrose, A. J., Francisco, A., Paulus, M. P., & Stein, M. B. (2015). Neural substrates of approach-avoidance conflict decision-making. *Human Brain Mapping*, 36(2), 449–462. doi:<https://doi.org/10.1002/hbm.22639>
- Bijleveld, E., Custers, R., & Aarts, H. (2010). Unconscious reward cues increase invested effort, but do not change speed-accuracy tradeoffs. *Cognition*, 115, 330–335. doi:<https://doi.org/10.1016/j.cognition.2009.12.012>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436. doi:<https://doi.org/10.1163/156856897X00357>
- Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 367–379. doi:<https://doi.org/10.3758/CABN.7.4.367>
- Carsten, T., Hoofs, V., Boehler, C. N., & Krebs, R. M. (2019). Are losses more effective than rewards in improving performance in a cognitive task? *Motivation Science*, 5(3), 257. doi:<https://doi.org/10.1037/mot0000117>
- Chajut, E., Mama, Y., Levy, L., & Algom, D. (2010). Avoiding the approach trap: A response bias theory of the emotional Stroop effect. *Journal of Experimental Psychology: Learning Memory and Cognition*, 36(6), 1567–1572. doi:<https://doi.org/10.1037/a0020710>
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, 97(3), 332–361. doi:<https://doi.org/10.1037/0033-295X.97.3.332>
- Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science*, 17(3), 222–227. doi:<https://doi.org/10.1111/j.1467-9280.2006.01689.x>
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222. doi:<https://doi.org/10.1126/science.1247003>
- Ert, E., & Erev, I. (2013). On the descriptive value of loss aversion in decisions under risk: Six clarifications. *Judgment and Decision Making*, 8(3), 214–235. doi:<https://doi.org/10.2139/ssm.1012022>
- Engelmann, J. B., & Pessoa, L. (2007). Motivation sharpens exogenous spatial attention, 1(S), 64–72. *Emotion*. doi:<https://doi.org/10.1037/1528-3542.7.3.668>
- Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General*, 130(4), 681–700. doi:<https://doi.org/10.1037/0096-3445.130.4.681>
- Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion*, 16(3), 355–379. doi:<https://doi.org/10.1080/02699930143000527>
- Gächter, S., Orzen, H., Renner, E., & Starmer, C. (2009). Are experimental economists prone to framing effects? A natural field experiment. *Journal of Economic Behavior & Organization*, 70, 443–446. doi:<https://doi.org/10.1016/j.jebo.2007.11.003>
- Grégoire, L., & Anderson, B. A. (2019). Semantic generalization of value-based attentional priority. *Learning & Memory*, 26, 460–464. doi:<https://doi.org/10.1101/lm.050336.119>
- Grégoire, L., Britton, M. K., & Anderson, B. A. (2020). Motivated suppression of value- and threat-modulated attentional capture. *Emotion*.
- Grégoire, L., & Greening, S. G. (2019). Opening the reconsolidation window using the mind's eye: Extinction training during reconsolidation disrupts fear memory expression following mental imagery reactivation. *Cognition*, 183, 277–281. doi:<https://doi.org/10.1016/j.cognition.2018.12.001>
- Grégoire, L., & Greening, S. G. (2020). Fear of the known: semantic generalisation of fear conditioning across languages in bilinguals. *Cognition & Emotion*, 34(2), 352–358. doi:<https://doi.org/10.1080/02699931.2019.1604319>
- Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2013). The musical Stroop effect: Opening a new avenue to research on automatisms. *Experimental Psychology*, 60(4), 269–278. doi:<https://doi.org/10.1027/1618-3169/a000197>
- Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2014). About the unidirectionality of interference: Insight from the musical Stroop effect. *Quarterly Journal of Experimental Psychology*, 67(11), 2071–2089. doi:<https://doi.org/10.1080/17470218.2014.896932>
- Grégoire, L., Perruchet, P., & Poulin-Charronnat, B. (2015). How does Stroop interference change with practice? A reappraisal from the musical Stroop paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, 41(2), 417–425. doi:<https://doi.org/10.1037/xlm0000044>
- Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in cognitive sciences*, 18(4), 194–202. doi:<https://doi.org/10.1016/j.tics.2014.01.003>
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291. doi:<https://doi.org/10.2307/1914185>
- Lejarraga, T., Schulte-Mecklenbeck, M., Pachur, T., & Hertwig, R. (2019). The attention-aversion gap: How allocation of attention relates to loss aversion. *Evolution and Human Behavior*, 40(5), 457–469. doi:<https://doi.org/10.1016/j.evolhumbehav.2019.05.008>
- Kim, A. J., & Anderson, B. A. (2019). Threat reduces value-driven but not salience-driven attentional capture. *Emotion*. doi:<https://doi.org/10.1037/emo0000599>
- Kim, H., & Anderson, B. A. (2020). How does the attention system learn from aversive outcomes? *Emotion*.
- Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 21(16), RC159 (1–5). doi:<https://doi.org/10.1523/jneurosci.21-16-j0002.2001>
- Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the Stroop task. *Cognition*, 117(3), 341–347. doi:<https://doi.org/10.1016/j.cognition.2010.08.018>
- Krypotos, A. M., Effting, M., Arnaudova, I., Kindt, M., & Beckers, T. (2014). Avoided by association: Acquisition, extinction, and

- renewal of avoidance tendencies toward conditioned fear stimuli. *Clinical Psychological Science*, 2(3), 336–343. doi:<https://doi.org/10.1177/2167702613503139>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, 109(2), 163–203. doi:<https://doi.org/10.1037/0033-2909.109.2.163>
- Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual Review of Neuroscience*, 24(1), 897–931. doi:<https://doi.org/10.1146/annurev.neuro.24.1.897>
- Müller, S., Rothermund, K., & Wentura, D. (2016). Relevance drives attention: Attentional bias for gain-and loss-related stimuli is driven by delayed disengagement. *The Quarterly Journal of Experimental Psychology*, 69(4), 752–763. doi:<https://doi.org/10.1080/17470218.2015.1049624>
- Murty, V. P., LaBar, K. S., & Alison Adcock, R. (2012). Threat of punishment motivates memory encoding via amygdala, not mid-brain, interactions with the medial temporal lobe. *Journal of Neuroscience* 32(26), 8969–8976. doi:<https://doi.org/10.1523/JNEUROSCI.0094-12.2012>
- Nissens, T., Failing, M., & Theeuwes, J. (2017). People look at the object they fear: Oculomotor capture by stimuli that signal threat. *Cognition and Emotion*, 31(8), 1707–1714. doi:<https://doi.org/10.1080/02699931.2016.1248905>
- Schindler, S., & Pfattheicher, S. (2017). The frame of the game: Loss-framing increases dishonest behavior. *Journal of Experimental Social Psychology*, 69, 172–177. doi:<https://doi.org/10.1016/j.jesp.2016.09.009>
- Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition and Emotion*, 29(4), 687–694. doi:<https://doi.org/10.1080/02699931.2014.924484>
- Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2017). The time course of attentional bias to cues of threat and safety. *Cognition and Emotion*, 31(5), 845–857. doi:<https://doi.org/10.1080/02699931.2016.1169998>
- Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, 14(2), 139–147. doi:<https://doi.org/10.1016/j.conb.2004.03.017>
- Stormark, K. M., Nordby, H., & Hugdahl, K. (1995). Attentional shifts to emotionally charged cues: Behavioural and ERP data. *Cognition and Emotion*, 9(5), 507–523. doi:<https://doi.org/10.1080/02699939508408978>
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643–662. doi:<https://doi.org/10.1037/h0054651>
- Thorndike, E. L. (1911). *Animal intelligence: Experimental studies* (The Animal Behavior Series). Lewiston, NY: Macmillan Press. doi:<https://doi.org/10.5962/bhl.title.55072>
- Van Damme, S., Lorenz, J., Eccleston, C., Koster, E. H. W., De Clercq, A., & Crombez, G. (2004). Fear-conditioned cues of impending pain facilitate attentional engagement. *Neurophysiologie Clinique*, 34(1), 33–39. doi:<https://doi.org/10.1016/j.neucli.2003.11.001>
- Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9(12), 585–594. doi:<https://doi.org/10.1016/j.tics.2005.10.011>
- Wentura, D., Müller, P., & Rothermund, K. (2014). Attentional capture by evaluative stimuli: Gain-and loss-connoting colors boost the additional-singleton effect. *Psychonomic Bulletin & Review*, 21(3), 701–707. doi:<https://doi.org/10.3758/s13423-013-0531-z>
- Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, 120(1), 3–24. doi:<https://doi.org/10.1037/0033-2909.120.1.3>
- Yechiam, E. (2019). Acceptable losses: The debatable origins of loss aversion. *Psychological Research*, 83(7), 1327–1339. doi:<https://doi.org/10.1007/s00426-018-1013-8>
- Yechiam, E., & Hochman, G. (2013a). Losses as modulators of attention: review and analysis of the unique effects of losses over gains. *Psychological Bulletin*, 139(2), 497. doi:<https://doi.org/10.1037/a0029383>
- Yechiam, E., & Hochman, G. (2013b). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive Psychology*, 66(2), 212–231. doi:<https://doi.org/10.1016/j.cogpsych.2012.12.001>
- Yiend, J., & Mathews, A. (2001). Anxiety and attention to threatening pictures. *The Quarterly Journal of Experimental Psychology Section A*, 54(3), 665–681. doi:<https://doi.org/10.1080/713755991>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.